

Aplicación móvil para el análisis y visualización De bases de datos bibliográficas

POR: CHRISTHIAN DAVID SANTANDER FIGUEROA

Trabajo de grado para optar por el título de Ingeniero de Sistemas

Universidad del Valle

Facultad de Ingeniería

Escuela de Ingeniería de Sistemas y Computación

Cali, Valle del Cauca, Colombia

Mayo 2015

APLICACIÓN MÓVIL PARA EL ANÁLISIS Y VISUALIZACIÓN

DE BASES DE DATOS BIBLIOGRÁFICAS

POR: CHRISTHIAN DAVID SANTANDER FIGUEROA

Trabajo de grado para optar el título de Ingeniería de Ingeniero de Sistemas

DIRECTOR: VÍCTOR A. BUCHELI

CODIRECTORA: BEATRIZ E. FLORIÁN



Facultad de Ingeniería

Escuela de Ingeniería de Sistemas y Computación

Cali, Valle del Cauca, Colombia

Mayo 2015

AGRADECIMIENTOS

A mis padres y mi familia por su apoyo incondicional, a mis profesores por guiarme durante estos años, a mis directores por sus aportes y guía. Gracias a los compañeros del proyecto de investigación Vigtech Univalle por los conceptos y discusiones que salieron de las reuniones de dicho proyecto.

LISTA DE FIGURAS

Figura 1 Aplicaciones Móviles	10
Figura 2 Flujo de la aplicación (Lógica)	16
Figura 3 Diagrama de despliegue (Física)	17
Figura 4 Menú VT-Mobile	20
Figura 5 Búsqueda VT-Mobile	21
Figura 6 Búsqueda Realizada	21
Figura 7 Menú de Analíticas	22
Figura 8 Tipos de gráficos y análisis	22
Figura 9 Barras Autores - documento	23
Figura 10 Circular tipo de documento	23
Figura 11 Circular Países - documento	24
Figura 12 Barras Países - documento	24
Figura 13 Barras Revistas - documento	25
Figura 14 Circular Instituciones - documento (Afiliaciones de los autores)	25
Figura 15 Barras Documentos por año	26
Figura 16 Clustering	26
Figura 17 Clustering expansión del grupo más grande	27
Figura 18 Clustering expansión segundo grupo	28
Figura 19 Clustering expansión tercer grupo	28
Figura 20 Lista de documentos	29
Figura 21 Lista de docentes	30
Figura 22 Pregunta 2	31
Figura 23 Pregunta 3	31
Figura 24 Pregunta 4	32
Figura 25 Pregunta 5	33
Figura 26 Pregunta 6	33
Figura 27 Pregunta 7	34
Figura 28 Pregunta 8	34
Figura 29 Pregunta 9	35
Figura 30 Pregunta 10	35

LISTA DE TABLAS

<i>Tabla 1 Resultados Objetivos Específicos</i>	<i>9</i>
<i>Tabla 2 Conceptos estudiados de minería de datos</i>	<i>13</i>
<i>Tabla 3 Análisis Comparativo</i>	<i>15</i>

CONTENIDO

Agradecimientos.....	3
Lista de Figuras.....	4
Lista de Tablas	5
Contenido	6
Resumen.....	7
CAPÍTULO 1 Introducción General	8
1.1 Problema y Motivación.....	8
1.2 Alcances y Contribuciones	8
1.3 Objetivos.....	9
1.3.1 Objetivo General.....	9
1.3.2 Objetivos Específicos.....	9
1.4 Resultados Obtenidos	9
1.5 Marco Teórico	10
1.5.1 Aplicaciones móviles.....	10
1.5.2 Revisión de conceptos de usabilidad	12
1.5.3 Scrum	12
1.5.4 Extracción de datos y agrupamiento.....	13
1.5.5 Antecedentes	14
CAPÍTULO 2 Proceso de Desarrollo	16
2.1 Arquitectura	16
2.2 Desarrollo	17
2.2.1 Aplicación móvil.....	17
2.2.2 Servidor.....	18
2.3 Metodología	18
2.4 Actividades de desarrollo	18
CAPÍTULO 3 Resultados y Conclusiones.....	20
3.1 Prueba Piloto.....	20
3.1.1 Prueba	20
3.1.2 Conclusiones de la prueba piloto	29
3.2 Prueba para la experiencia de usuario	29
3.2.2 Conclusiones prueba de experiencia de usuario	36
3.3 Conclusiones	36
3.4 Trabajo Futuro	36
Referencias	37

RESUMEN

Este documento de trabajo de grado presenta un prototipo de aplicación móvil (app), VT Mobile, que apoya la vigilancia científico tecnológica en la Universidad del Valle, parte del proyecto de investigación de convocatoria interna *Hacia el estudio de estrategias computacionales para potenciar la vigilancia tecnológica; un enfoque interdisciplinar en ingeniería de sistemas y computación*. El prototipo produce un conjunto de visualizaciones e indicadores a partir de la extracción de información de un conjunto de datos de Scopus. Para producir dichas visualizaciones e indicadores los datos son enriquecidos siguiendo las actividades de: búsqueda, extracción, agrupamiento, análisis de datos y visualización de datos. Esta herramienta permite hacer análisis sobre la producción científica de la Escuela de Ingeniería de Sistemas y Computación (EISC). Este documento presenta una descripción de las bases conceptuales que fundamentan el desarrollo de la herramienta VT Mobile, así como una descripción de su arquitectura, desarrollo y funcionalidad, finalmente se presenta un estudio piloto para validar la integridad de los datos producidos y un estudio sobre los profesores de la EISC para analizar la experiencia de usuario que provee la aplicación y la posición de los docentes frente al desarrollo de esta herramienta.

CAPÍTULO 1

INTRODUCCIÓN GENERAL

Una de las actividades a las que se dedican las comunidades académicas es a la investigación. Los procesos de investigación requieren información y un seguimiento amplio de los cambios del campo estudio y las nuevas publicaciones; las bases de datos bibliográficas proveen una búsqueda de todo tipo de artículos, patentes, etc. Algunas incluso con herramientas de análisis para organizar los autores más referenciados, los que han realizado más contribuciones y otros datos bibliográficos. Este tipo de información se encuentra en bases de datos tales como Scopus y requiere ser analizada cada vez. La Vigilancia tecnológica es un proceso organizado que capta información sobre una organización para analizarla o difundirla y convertirla en conocimiento que pueda apoyar la toma de decisiones, las comunidades científicas son responsables de la producción de nuevo conocimiento y este tipo de prácticas se hace necesario para monitorear esta producción y obtener información que permita beneficiar a la comunidad y a la producción de conocimiento.

1.1 Problema y Motivación

La Universidad del Valle aprobó el proyecto *Hacia el estudio de estrategias computacionales para potenciar la vigilancia tecnológica; un enfoque interdisciplinar en ingeniería de sistemas y computación* cuyo objetivo es desarrollar un prototipo de plataforma computacional para el análisis de datos como herramienta de soporte orientada a la vigilancia tecnológica del sistema regional de ciencia y tecnología del valle del cauca, Vigtech Univalle. Permitir la observación de la producción de conocimiento y el monitoreo de dicha producción. Las prácticas de vigilancia tecnológica buscan el estudio de comunidades científicas, dado que son estas las que permiten el desarrollo del conocimiento. Planeación, búsqueda y captación, análisis y organización, inteligencia y por ultimo comunicación son las fases del ciclo de vigilancia tecnológica de acuerdo a (Sánchez Torres, J. M., & Palop Marro, F. (2008)) Por tanto, las comunidades científicas son el foco de estudio para la vigilancia tecnológica. Para poder realizar estos procesos eficientemente se requieren herramientas como Vigtech Univalle pero para poder realizar el monitoreo de estas actividades y obtener los resultados esperados se necesita que la aplicación sea robusta y simple de usar por lo cual se hace necesario una herramienta móvil que pueda presentar la información de forma simple y organizada dándole un valor agregado a los datos y posibilidad de acoplarse a una plataforma más grande (Vigtech Univalle).

1.2 Alcances y Contribuciones

Se implementó una aplicación móvil híbrida, VT-Mobile, para el apoyo de la vigilancia científico-tecnológica en la Universidad del Valle. El prototipo produce un conjunto de visualizaciones e indicadores a partir de la extracción de información de un conjunto de datos de Scopus. Para producir dichas visualizaciones e indicadores los datos son

enriquecidos siguiendo las actividades de: búsqueda, extracción, agrupamiento, análisis de datos y visualización. Esta herramienta permite hacer análisis sobre la producción científica de la Escuela de Ingeniería de Sistemas y Computación. Se pueden realizar búsquedas sobre la sintaxis de Scopus, obtener indicadores y visualizaciones que despliegan la información de forma sencilla y rápida.

1.3 Objetivos

1.3.1 Objetivo General

Desarrollar una aplicación móvil para apoyar los procesos de investigación académica a través de algoritmos de agrupamiento y Academic analytics, extraer, agregar, ordenar y visualizar información proveniente de bases de datos bibliográficas de temas y autores.

1.3.2 Objetivos Específicos

- Definir las técnicas para la extracción de datos y agrupamiento.
- Definir las tecnologías para implementar la aplicación.
- Desarrollar el módulo de extracción de datos.
- Desarrollar el modulo para la agrupación de temas o documentos usando algoritmos de agrupación.
- Desarrollar el modulo principal para el análisis y despliegue grafico tomando en cuenta los principios de usabilidad como por ejemplo (Nielsen & Molich, 1990).

1.4 Resultados Obtenidos

Tabla 1 Resultados Objetivos Específicos

Objetivos	Resultados	Desarrollo
Definir las técnicas para la extracción de datos y agrupamiento	Se comprende que poseer una gran cantidad de datos no provee por sí misma información que permita decisiones o análisis sobre el caso de estudio específico, esta información debe ser interpretada de alguna forma para ello se pre procesa obteniendo conceptos importantes y relaciones que no pueden ser detectadas con facilidad dada la cantidad abrumadora de datos de donde provienen.	Se estudiaron conceptos básicos de Data Mining para la limpieza de información y descubrimiento del conocimiento, se tomaron en cuenta conceptos como: árboles de decisión, reglas de clasificación, reglas de asociación, regresión lineal, bayes y clustering. Se determinó usar clustering para la agrupación de documentos similares y reglas simples para la selección de información.
Definir las tecnologías para implementar la aplicación.	Se entienden las diversas alternativas para construir la aplicación cada una con sus puntos a favor y en contra. Se elige finalmente seguir el camino de las aplicaciones híbridas dado que se ajustan a las necesidades de la aplicación y proveen beneficios ya que permiten el uso de herramientas ya probadas y conocidas sin ahondar en las especificidades de cada dispositivo las cuales pueden variar en gran medida, exigiendo diferentes desarrollos para otras plataformas.	Hay diferentes tipos de aplicaciones móviles: Web, híbridas y nativas cada una tiene ciertas ventajas se tomó en cuenta cada una para luego determinar que la aplicación sería híbrida. Una aplicación que no necesitaría una conexión constante a internet para revisar el análisis anterior, con acceso a las características del dispositivo, más sencillo de portar a otras plataformas y con tecnologías ya conocidas del desarrollo web.

Desarrollar el módulo de extracción de datos	Se construye un módulo capaz de obtener los datos requeridos y de extraer las características necesarias para el proceso tomando en cuenta las técnicas estudiadas.	Se implementó un módulo que se comunica directamente con Elsevier y descarga información semi-estructurada, también se lleva a cabo la extracción de características: vectores de características documentales y funciones de Distancia propias para este tipo de datos por ejemplo jaccard. Para identificar información: útil agrupamiento de documentos.
Desarrollo del módulo de agrupamiento	Se tiene un módulo que toma a cada documento como un vector de características y realiza clustering para agrupar los documentos basados en las palabras claves que comparten.	Se utiliza clustering sobre los datos pre procesados para agrupar los documentos por características similares y luego producir una visualización de esto.
Desarrollo del módulo principal para análisis y despliegue gráfico de la aplicación tomando en cuenta la usabilidad.	Se estudió la librería para el desarrollo web mobile JQuery mobile y las librerías para la visualización. D3 ofrece muchas posibilidades para representar datos de todo tipo de manera fluida y agradable por lo cual se implementó un gráfico de red agrupada, todo esto fue empaquetado usando phonegap e instalado como una aplicación.	El análisis de redes sociales, toma elementos del álgebra de matrices al igual que de la teoría de grafos para construir desde un conjunto delimitado de actores vinculados entre sí, una representación de las relaciones existentes, dichos actores pueden vincularse de diferentes modos, adyacencia, afiliación o atributos. De esta forma es posible modelar, medir y visualizar las estructuras sociales, identificando patrones, se implementó entonces en la aplicación la visualización de una red agrupada.

1.5 Marco Teórico

1.5.1 Aplicaciones móviles

Para el desarrollo de la aplicación se investigaron diferentes tecnologías y conceptos, hay diferentes tipos de desarrollo móvil mencionados en (Quintero, C. A., Florian-Gaviria, B., & Pabon, O. S. 2014). A continuación se da una descripción general y se explica cuál fue la elección para este desarrollo.

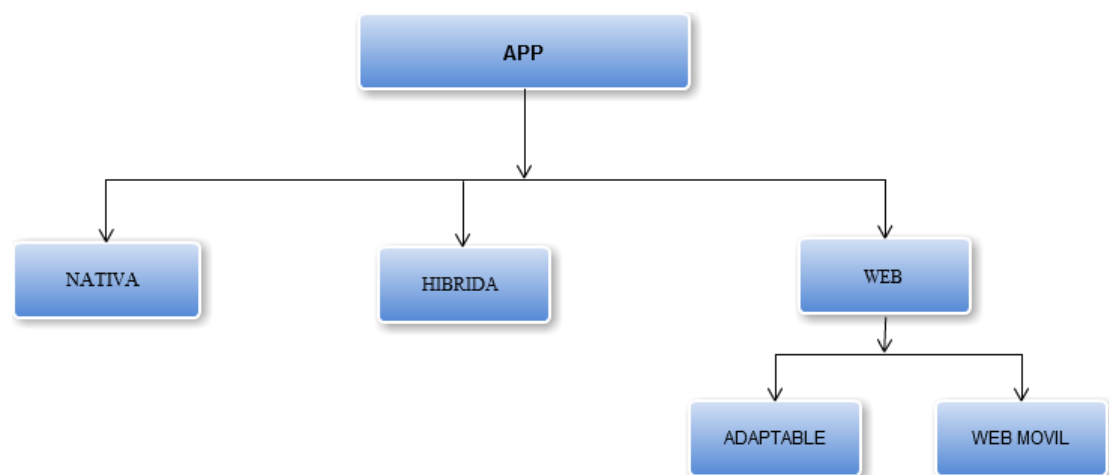


Figura 1 Aplicaciones Móviles

1.5.1.1 *Aplicaciones Nativas*

Estas son desarrolladas dentro del lenguaje específico del dispositivo, tienen fácil acceso a todas las capacidades de este y suelen ser más eficientes que las de otros tipos, una aplicación nativa cuenta con grandes ventajas, no se pueden superar en rendimiento pues están hechas específicamente para el dispositivo pero esto también es un problema. Las aplicaciones nativas requieren el aprendizaje del lenguaje de programación que usan las aplicaciones de ese dispositivo en particular, esto aumenta la cantidad de versiones necesarias para que funcione en distintas plataformas y también sufren la falta de librerías en este caso librerías para la visualización de datos.

1.5.1.2 *Aplicaciones Web*

Dentro de esta categoría se encuentran dos tipos, adaptables y web móvil, ambas consisten de usar tecnologías web para crear la aplicación, crear un sitio web adaptable quiere decir que la página tratará de ajustarse al tamaño del dispositivo para desplegar su contenido, un ejemplo de esto son las paginas responsive, mientras que Web móvil se refiere a sitios contruidos específicamente para dispositivos diferentes a un computador y se usan librerías Css y java script para simular la forma de desplazarse particular de los dispositivos móviles. Las páginas adaptables pueden tener problemas al desplegar el mismo contenido en diferentes pantallas, es una sola página que cambia dependiendo del dispositivo pero no siempre se obtienen los mejores resultados. Los sitios web móviles al ser contruidos específicamente para móviles se desempeñan bastante bien pero son versiones extra de una página web así que se requiere mantenimiento a dos páginas distintas, la principal y la versión móvil. Ambas alternativas permiten hacer uso de librerías hechas por terceros pero también tienen desventajas, para las aplicaciones web es difícil usar las características del dispositivo (almacenamiento, cámara, etc.) para acceder a estos tipos de aplicación es necesario tener una conexión al servidor en el cual se encuentra el sitio.

1.5.1.3 *Aplicaciones Híbridas*

Estas fueron las elegidas para desarrollar la aplicación, son una mezcla entre las aplicaciones web y nativas, se trabaja con las tecnologías web ya conocidas pero la aplicación se instala en el celular como si fuera nativa, existen plugins para acceder a las características del dispositivo y no es necesaria una conexión a internet para poder ejecutar la aplicación. Las aplicaciones híbridas usan HTML, CSS y Java Script pero se inserta código nativo para que se cree una webview (visor nativo de páginas web) que despliegue todo como si fuera el navegador, se puede usar la aplicación sin conexión, herramientas de terceros en HTML, CSS y Java Script y acceder a las funciones básicas del dispositivo, la desventaja de este tipo de aplicaciones es el rendimiento pues al no ser una aplicación nativa su rendimiento es menor y más limitado que una que sí lo es. Se escogió este tipo de aplicación para poder trabajar con tecnologías ya conocidas, hacer uso de librerías de visualización y tener la posibilidad de usar VT-Mobile sin necesidad de conexión.

1.5.2 Revisión de conceptos de usabilidad

Se revisaron heurísticas de usabilidad y se tomaron en cuenta como lineamientos los siguientes conceptos:

Optimizar la navegación

Evitar que el usuario se pierda en la aplicación, delimitar claramente los elementos de navegación pero sin saturar el espacio.

Comprobar gráficos y colores

No utilizar medidas absolutas para las dimensiones de objetos, tener en cuenta el tamaño de los elementos gráficos y no textuales, asegurarse del contraste entre los elementos para no saturar la visión o perder legibilidad.

Hacer en pequeño

Reducir el tamaño de las páginas y hojas de estilo, que no se necesiten barras de desplazamiento, minimizar para gastar menos recursos.

Economizar el uso de la red

No usar actualizaciones ni redirecciones automáticas.

Facilitar la entrada de datos

Minimizar el uso del teclado ya que el uso de esto en los dispositivos móviles puede ser tedioso, establecer valores por defecto, etiquetar generosamente los formularios, reducir la introducción de texto a lo necesario.

Pensar en los usuarios móviles

Limitar el contenido a lo que el usuario necesita, títulos descriptivos y lenguaje sencillo y claro.

Ninguna heurística es perfecta ni existe un consenso general sobre usabilidad pero se tomaron en cuenta ideas usadas en el desarrollo web anteriormente para formar los conceptos para el desarrollo de la aplicación.

1.5.3 Scrum

Es una metodología ágil con iteraciones cada 15 o 30 días dependiendo de lo pactado al inicio del proyecto y con reuniones diarias, es un proceso iterativo e incremental en el cual se plantea una lista de tareas de todo el producto (product backlog) y luego se define una para cada sprint (periodo de 15 o 30 días) los equipos toman las tareas y las implementan, al final del sprint se hacen 2 reuniones en las cuales se revisa lo producido, lo terminado y lo incompleto se formula un demo o entregable que muestra lo que se completó y se revisa y rediseña si es necesario el backlog, se plantea la siguiente lista y el proyecto pasa por otro sprint y así sucesivamente, con las reuniones diarias y a final de cada sprint se identifican cambios, problemas y avances y el proyecto avanza y se mejora con cada iteración. Todo lo anterior basado en Sutherland, J., & Schwaber, K. (2007)

1.5.4 Extracción de datos y agrupamiento

1.5.4.1 Minería de datos

“La minería de datos puede definirse como un proceso de descubrimiento de nuevas y significativas relaciones, patrones y tendencias al examinar grandes cantidades de datos”
López, C. P. (2007).

Tabla 2 Conceptos estudiados de minería de datos

Concepto	Características	Ventajas	Desventajas
Arboles de decisión	Se forman árboles a partir de los valores que puedan tomar los atributos en cada instancia.	No existe ambigüedad.	La Información incompleta genera grandes problemas, sobre entrenamiento.
Reglas de clasificación y asociación	se descubren patrones o reglas que permiten predecir un atributo o los demás atributos	Menos reglas que las extraídas directamente de un árbol de decisión.	Limitadas, dependiendo de su tipo secuenciales.
Regresión lineal	Modela una relación entre variables independientes y una variable dependiente.	preciso	No todos los problemas pueden ser modelados de forma lineal.
Bayes ingenuo	Es un clasificador probabilístico, clasifica en base a probabilidades.	Es fácil de implementar y obtiene buenos resultados.	Sensibilidad en los resultados dependiendo de las distribuciones.
Agrupamiento clustering	A partir de una matriz de distancias se forman n grupos diferentes con elementos similares.	Genera relaciones o jerarquías.	Se necesita el número de grupos o una medida de éxito para evaluar los grupos.

Para procesar la información existen múltiples alternativas algunas de ellas contenidas dentro de la tabla anterior de las cuales se escogieron dos para usarse en el proyecto, Reglas y clustering, dado que este proyecto es una primera versión es conveniente empezar con algo simple por esto se escogió como alternativa a las reglas de clasificación y asociación pues son bastante simples y proveen buenos resultados, se usaron para pre procesar la información y obtener una entrada con la que el clustering pueda trabajar, escogido para poder encontrar grupos de documentos por temas.

1.5.4.1.1 Reglas de Asociación

Existen para descubrir patrones o información escondida dentro de los datos iniciales como por ejemplo un supermercado que mide la frecuencia de compras de ciertos objetos en periodos de tiempo para luego descubrir que objetos se compran juntos y pueden ser usados en alguna estrategia de ventas, en el caso de la aplicación se toman las palabras claves y sus apariciones en distintos campos de los documentos para determinar al final con clustering el grupo de temas al que pertenecen.

1.5.4.1.2 Clustering

Es la tarea de agrupar elementos en grupos que compartan cierta similitud en este caso partiendo de la entrada que se consigue a través de las palabras claves y los campos se puede generar una matriz de distancias con la cual se puede juzgar que tan parecidos son ciertos elementos con lo cual el algoritmo de clustering ordenara todos en sus grupos respectivos los cuales serán representados gráficamente o mostrados como los diferentes temas que tratan los documentos de la consulta.

1.5.5 Antecedentes

Se estudiaron dos herramientas que comparten algunas características con la aplicación, Vigtech (Bucheli, V.A 2007) fue base fundamental para la creación de VT-Mobile mientras que Datafinder fue una guía respecto a cómo se presentan datos o visualizaciones en dispositivos móviles, se escogió Datafinder para la comparación dado que no se encontró ninguna aplicación móvil enfocada a la vigilancia tecnológica de la forma en que lo está VT-Mobile, a continuación se da una descripción corta de cada una y se hace una comparación.:

1.5.5.1 *Análisis comparativo de herramientas con características similares.*

1.5.5.1.1 Data world bank (Datafinder app)

Datafinder es una aplicación desarrollada por el banco mundial para dispositivos móviles, capaz de ordenar y graficar, toda la información proviene de la API del banco mundial la cual está repleta de indicadores internacionales de distinto tipo, es una herramienta completa y bastante útil pero a diferencia de la aplicación presentada en este documento, Datafinder no agrega valor a la información ni se realizan análisis sobre esta. Es una aplicación enfocada al despliegue de datos estadísticos y se toma como referencia por ser una aplicación móvil que ordena y presenta información.

1.5.5.1.2 Vigtech

Vigtech (Bucheli, V.A 2007) es una herramienta para el apoyo a las prácticas de vigilancia tecnológica enfocada en el ámbito científico tecnológico. La herramienta permite encontrar relaciones cognitivas y sociales en un conjunto de documentos extraídos de una base referencial en este caso SCOPUS a través de un crawler. Vigtech soporta las actividades de obtención de información de documentos científicos, extracción de metadatos, cálculo de estadísticas descriptivas, análisis de redes sociales, análisis de redes de palabras claves y visualización. La herramienta Vigtech es una herramienta web que se ajusta al ciclo de la vigilancia tecnológica (fases de una aplicación de inteligencia competitiva) Sánchez Torres, J. M., & Palop Marro, F. (2008) y por lo tanto se toma como base para el desarrollo de la aplicación móvil.

Tabla 3 Análisis Comparativo (Herramientas con características similares.)

Información general de la app				Fases de vigilancia tecnológica (aplicación de inteligencia competitiva)					
Producto	Descripción	Información que utiliza	Móvil	Búsqueda y Adquisición	Planeación	Búsqueda y Captación	Análisis e información	Inteligencia	Comunicación
Vigtech	Es una herramienta para vigilancia tecnológica	Semi-estructurada	No	Crawler	No	Automatizada	Descriptivos, Multivariados.	Cienciometría, Bibliometría, Análisis de redes sociales, minería de datos.	No
Datafinder WorldBank	Es una aplicación móvil para el despliegue de datos estadísticos	Estructurada	Si	WorldBank API	No	Automatizada	Descriptivos.	No	No
VT-Mobile	Es una aplicación móvil para la vigilancia tecnológica	Semi-estructurada	Si	Elsevier API	No	Automatizada	Descriptivos, Multivariados.	Cienciometría, Bibliometría, minería de datos.	No

La aplicación toma a Vigtech (Bucheli, V.A 2007) como punto de partida para recrear algunas de sus funciones pero en plataformas móviles en donde no se encontró ningún tipo de herramienta con propósito similar, Data finder comparte algunas características pero a diferencia de ésta VT-Mobile toma en cuenta el proceso de la vigilancia tecnológica y no solo indicadores descriptivos. VT-Mobile hace parte de un proyecto más grande para desarrollar Vigtech Univalle así que toma características de la herramienta Vigtech y las recrea dentro de dispositivos móviles. Para VT-Mobile se toma en cuenta Vigtech y se migran algunos procesos de web a móvil, mientras que Datafinder se toma solo como referencia de aplicación móvil para el despliegue de datos.

PROCESO DE DESARROLLO

La aplicación fue desarrollada con base en las funcionalidades de Vigtech y pensando en su interacción futura con la plataforma Vigtech Univalle de la cual hace parte. VT-Mobile es entonces una aplicación diseñada para brindar posibilidades móviles a la plataforma Vigtech Univalle y con su objetivo propio de tener utilidad dentro de la EISC analizando información sobre las publicaciones de cada autor dentro de la misma. Así entonces se diseñó con un enfoque en 3 de las fases del ciclo de vigilancia tecnológica e inteligencia competitiva: 1) Búsqueda y captación, 2) Análisis e 3) Inteligencia. La arquitectura de la aplicación se presenta de forma lógica y física en las figuras 2 y 3.

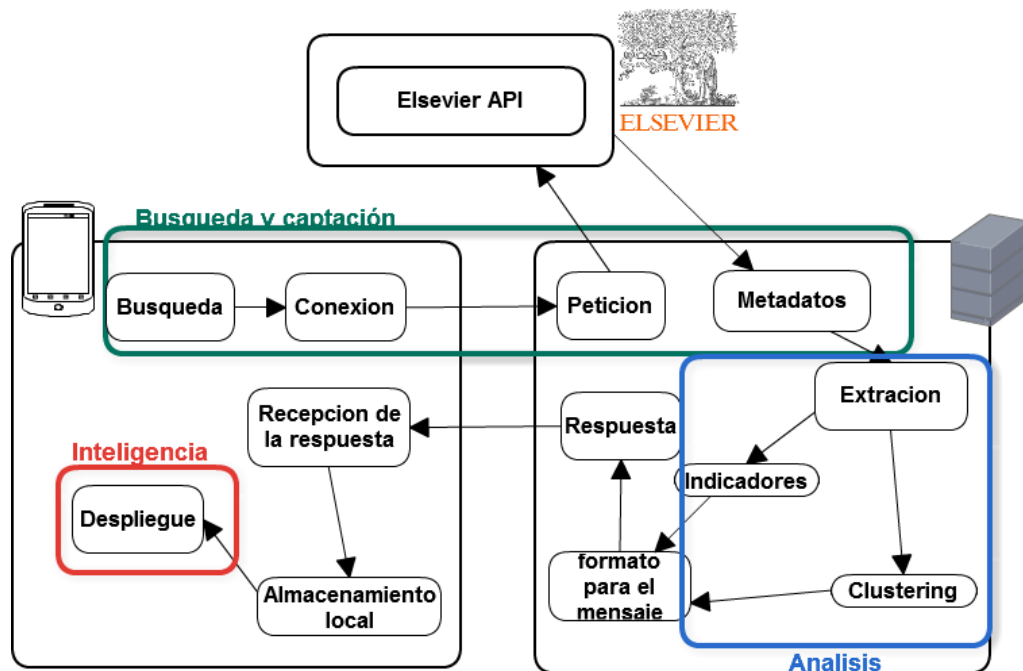


Figura 2 Flujo de la aplicación (Lógica)

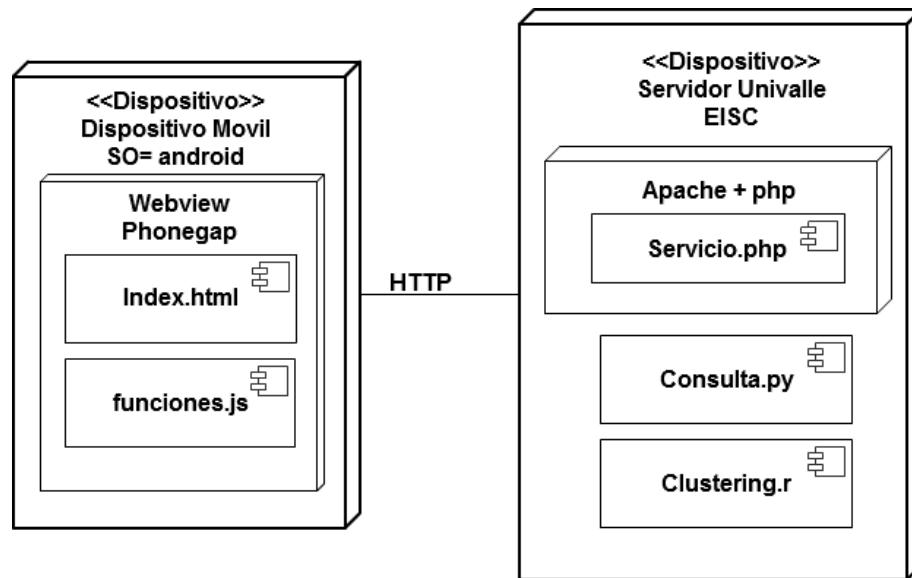


Figura 3 Diagrama de despliegue (Física)

Tal y como se puede apreciar en el diagrama de despliegue (figura 3), la aplicación se encuentra dividida en 2 partes: el dispositivo móvil que presenta todas las opciones y resultados de la aplicación para los diferentes usuarios y el servidor encargado del procesamiento. Los dispositivos móviles en su mayoría no poseen capacidades superiores de procesamiento por lo cual realizar los análisis directamente en el celular ocasionaría retrasos y problemas que irían en contra de la experiencia del usuario, haciendo imposible el uso en algunos dispositivos. El dispositivo móvil actúa como una interfaz que despliega todas las opciones y resultados; el servidor contiene la lógica de la aplicación, obtiene la información la procesa y envía una respuesta.

2.2 Desarrollo

Para desarrollar la aplicación se utilizó una arquitectura cliente - servidor. El dispositivo móvil actúa como cliente y consume un servicio web en el servidor, a continuación se describe la implementación:

2.2.1 Aplicación móvil

Después de estudiar las distintas oportunidades de desarrollo móvil y elegir que VT-Mobile fuera una aplicación híbrida, esta se implementó como tal usando librerías de JavaScript como JQuery Mobile¹, D3² y Highcharts³ como si fuera un sitio web móvil que luego sería empaquetado por Phonegap para su instalación (archivo apk) en un dispositivo móvil, en este caso dispositivos Android. La aplicación se diseñó enteramente para la visualización así que dentro de sus funciones esta realizar una petición y guardar la respuesta para el despliegue de indicadores y visualización grafica de los datos de la búsqueda y del resultado del clustering. También se usa la respuesta para generar una lista de los documentos estudiados para acceder a su ubicación en SCOPUS si es necesario.

2.2.2 Servidor

El servidor es responsable de crear una respuesta en formato Json que esté construida específicamente para ser visualizada con alguna de las librerías dentro de la aplicación móvil para obtener esto primero debe hacer una petición a Scopus y descargar la respuesta que en este caso es un XML, todo esto está dentro de un servicio web programado en Php que llama a Python para la petición y descarga del XML, luego es capaz de extraer la información importante de esta respuesta de la cual ciertos indicadores se ponen en el formato adecuado y el resto de la información extraída es sometida a otro proceso realizado en R el cual haciendo uso de la misma metodología explicada en (Bucheli, V. A. 2007) agrupa los documentos basados en características similares, estos grupos son especificados con el nombre del documento y enviados como parte del mensaje para la aplicación.

2.3 Metodología

Para el desarrollo de VT-Mobile se tomó en cuenta la metodología de desarrollo ágil Scrum y se tomaron algunas bases de esta para desarrollar el proyecto enfatizando en las reuniones frecuentes y la comunicación, se realizaron reuniones frecuentes y entregas constantes con un backlog para el proyecto basado en los objetivos y uno por cada reunión (sprint) las iteraciones se mantuvieron en periodos de 1 semana en las cuales al final se discutió el estado del proyecto los avances, retrasos, nuevas tareas y se entregaba un prototipo o pequeño avance respecto a los módulos de la aplicación.

2.4 Actividades de desarrollo

En resumen la lista de actividades para cumplir con los requerimientos de esta aplicación es la siguiente:

- i. Definir las técnicas para la extracción de datos y agrupamiento. Se estudiaron conceptos de Data Mining y se decidió que se utilizaría clustering para el agrupamiento, se extraen características similares con reglas simples de cada documento y luego se agrupan con clustering. Se comprende que poseer una gran cantidad de datos no provee por sí misma información que permita decisiones o análisis sobre el caso de estudio específico, esta información debe ser interpretada de alguna forma para ello se pre procesa obteniendo conceptos importantes y relaciones que no pueden ser detectadas con facilidad dada la cantidad abrumadora de datos de donde provienen.
- ii. Definir las tecnologías (móviles) para la implementación del prototipo. se consideran las alternativas para construir la aplicación (web, nativa, híbrida) cada una con sus puntos a favor y en contra. se elige finalmente seguir el camino de las aplicaciones híbridas dado que se ajustan a las necesidades de la aplicación y permiten el uso de herramientas y librerías ya construidas sin ahondar en las especificidades de cada dispositivo las cuales pueden variar en gran medida, exigiendo diferentes desarrollos para otras plataformas.

- iii. Desarrollar el módulo de extracción de datos. Para el desarrollo de este módulo se utilizaron los lenguajes Php y Python, se construyó un módulo capaz de obtener los datos requeridos y de extraer las características necesarias para el proceso tomando en cuenta las técnicas estudiadas.
- iv. Desarrollar el módulo de agrupamiento. Se crea un módulo en R Gentleman, R., Ihaka, R., & Bates, D. (2009) que toma a cada documento como un vector de características y realiza clustering para agrupar los documentos basado en las palabras claves que comparten, todo esto usando índice jaccard y basado en el trabajo descrito en Bucheli, V. A. (2007).
- v. Desarrollar el módulo principal para análisis y despliegue gráfico de la aplicación. Para desarrollar este módulo que compete a toda la aplicación dentro del dispositivo móvil se estudiaron distintas librerías. la librería para el desarrollo web móvil JQuery Mobile y las librerías para la visualización. D3 y Highcharts fueron utilizadas para construir este módulo, se implementó una visualización de red agrupada y varios otros gráficos estándar (circular, barras). todo fue empaquetado usando Phonegap e instalado como una aplicación que abre un webview en el cual se ejecuta todo el desarrollo realizado en HTML, CSS y Java Script.
- vi. Realizar pruebas sobre el prototipo. Se realizó una búsqueda sobre un autor específico para mostrar indicadores, gráficos y la visualización del agrupamiento, todo esto se muestra más adelante en la sección de resultados.

CAPÍTULO 3

RESULTADOS Y CONCLUSIONES

3.1 Prueba Piloto

3.1.1 Prueba

Se realizó una prueba piloto en donde se muestra la aplicación en su estado funcional durante el proceso de búsqueda de la producción científica de un autor específico, en este caso Fabio A. Gonzales O. será el autor elegido para comprobar el funcionamiento de la aplicación y se mostrara en detalle las funcionalidades de VT-Mobile.

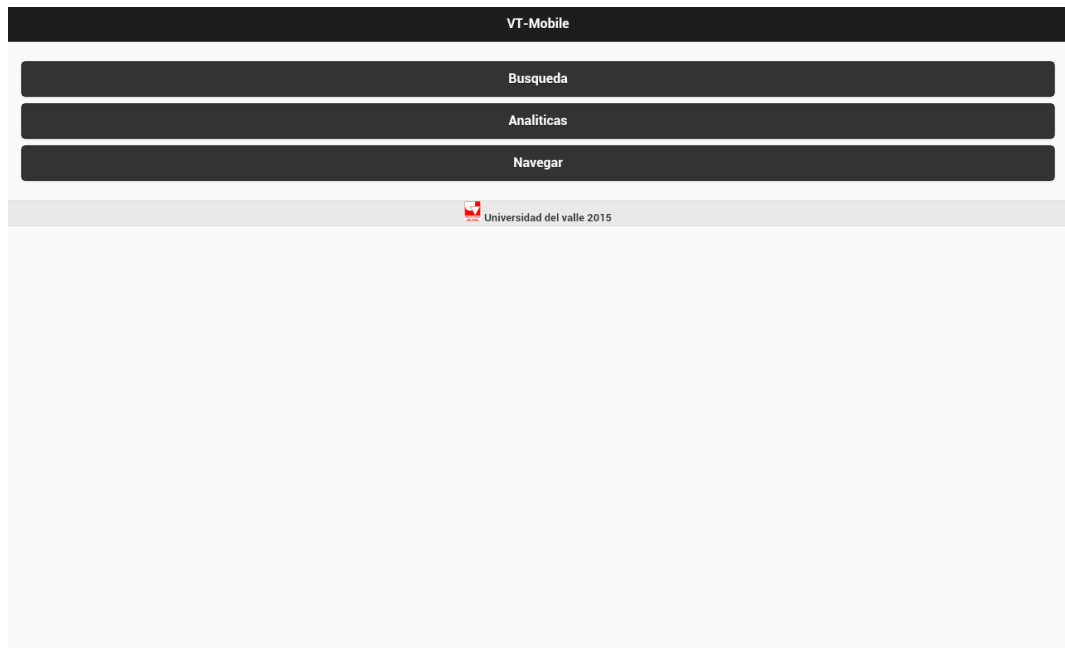


Figura 4 Menú VT-Mobile

En la figura 4 se ve el menú actual de VT-Mobile con las opciones de búsqueda, analíticas y navegar. Estas corresponden a la consulta, los resultados, visualizaciones y una lista de los documentos con su enlace al lugar en el que se encuentran dentro de Scopus.

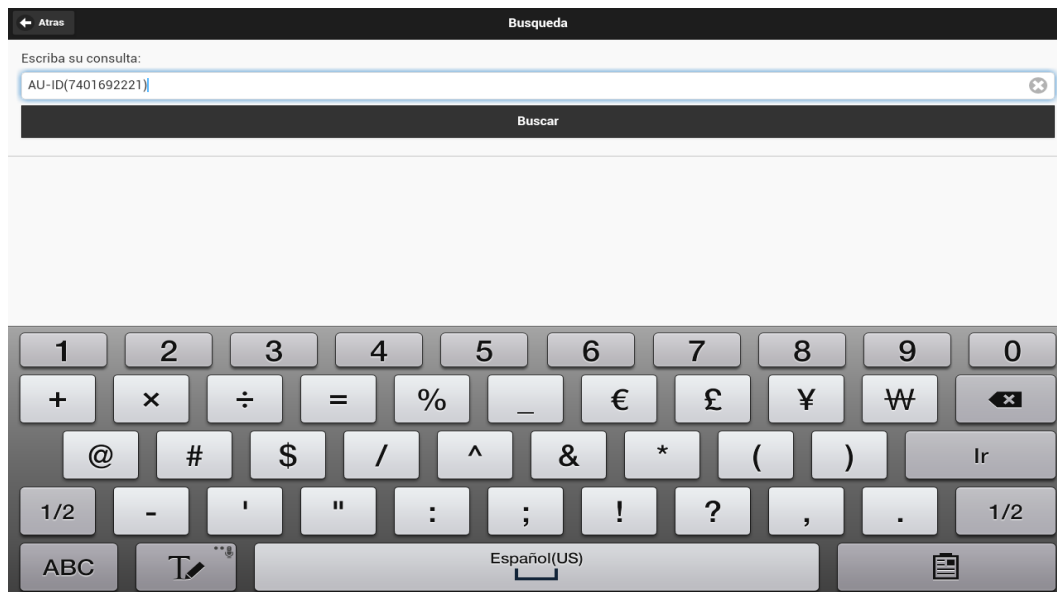


Figura 5 Búsqueda VT-Mobile

La búsqueda es bastante sencilla pero se puede utilizar la sintaxis de Scopus para obtener resultados más específicos como en este caso, en la figura 5 se realiza la consulta AU-ID(7401692221) que corresponde a la búsqueda por ID y el id corresponde a Fabio A. Gonzales O.

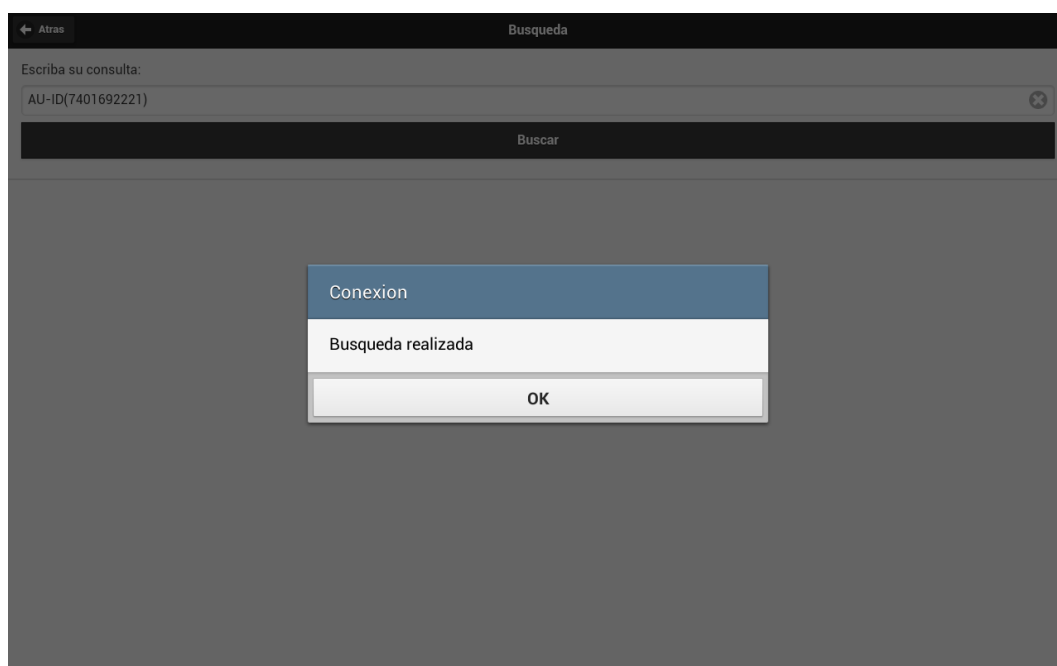


Figura 6 Búsqueda Realizada

←

Atras

Analíticas

Datos generales:

Consulta:

AU-ID(7401692221)

H-Index:

8

Promedio autores por documento:

2

Promedio de citas por documento:

12

Total de documentos:

52

Total de citas:

629

Total de autores:

80

Total de revistas:

22

Tipo de grafico:

Bar Chart

Autores

Tipo de documento

Países

Revistas

Afiliaciones

Documentos por año

Tipo de analisis:

Figura 7 Menú de Analíticas

En la figura 7 se muestra la mayor parte del menú de analíticas en las cuales se encuentran los indicadores, tipos de gráfico y tipos de análisis los cuales se muestran en la figura 8, VT-Mobile actualmente soporta clustering.

AU-ID(7401692221)

8

Promedio autores por documento:

2

Promedio de citas por documento:

12

Total de documentos:

52

Total de citas:

629

Total de autores:

80

Total de revistas:

22

Tipo de grafico:

Bar Chart

Autores

Tipo de documento

Países

Revistas

Afiliaciones

Documentos por año

Tipo de analisis:

Clustering

Analisis de Redes sociales

Reduccion de Dimensionalidad

Figura 8 Tipos de gráficos y análisis

Dentro del menú de analíticas se encuentran en indicadores, la consulta que se realizó, el H-Índex correspondiente que nos da información acerca del impacto de los artículos encontrados, las citaciones totales, el total de documentos encontrados, revistas, autores, y el promedio de citas y de autores por documento. Dentro de los tipos de grafico esta la selección entre grafico de barras y circular para los autores, documentos, países, revistas, afiliaciones y documentos por año. Cada uno dentro de las figuras siguientes:

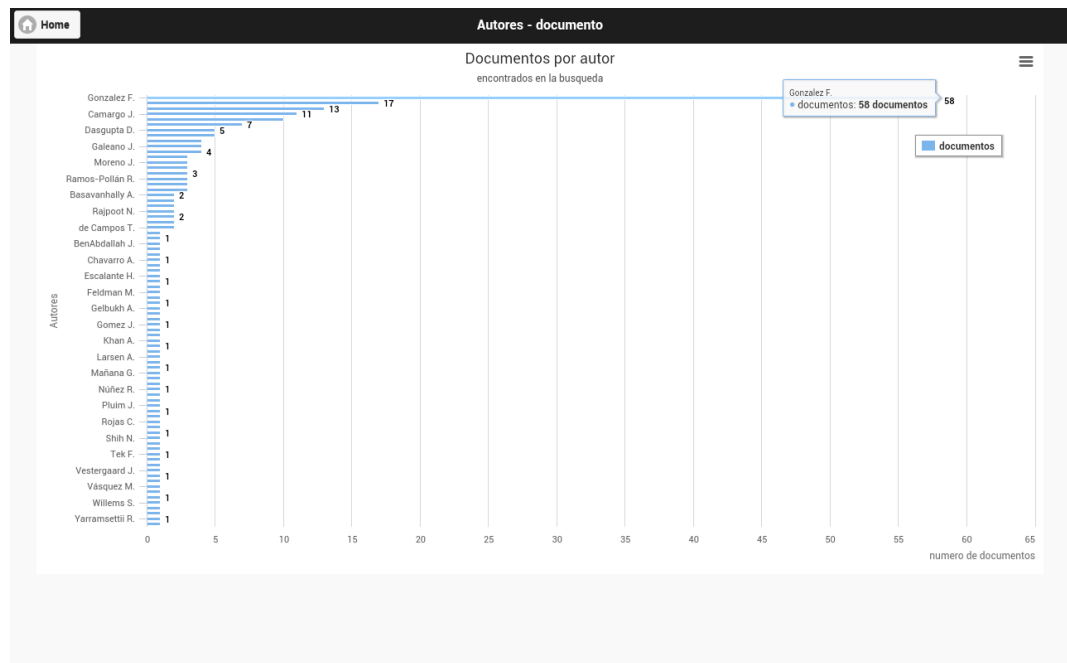


Figura 9 Barras Autores - documento

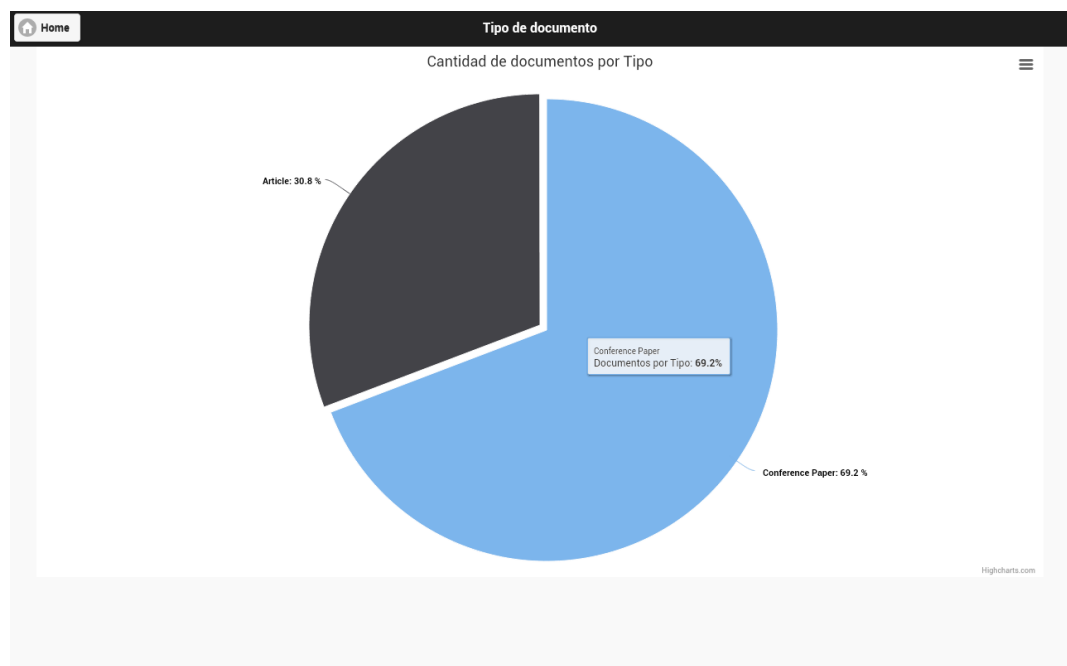


Figura 10 Circular tipo de documento

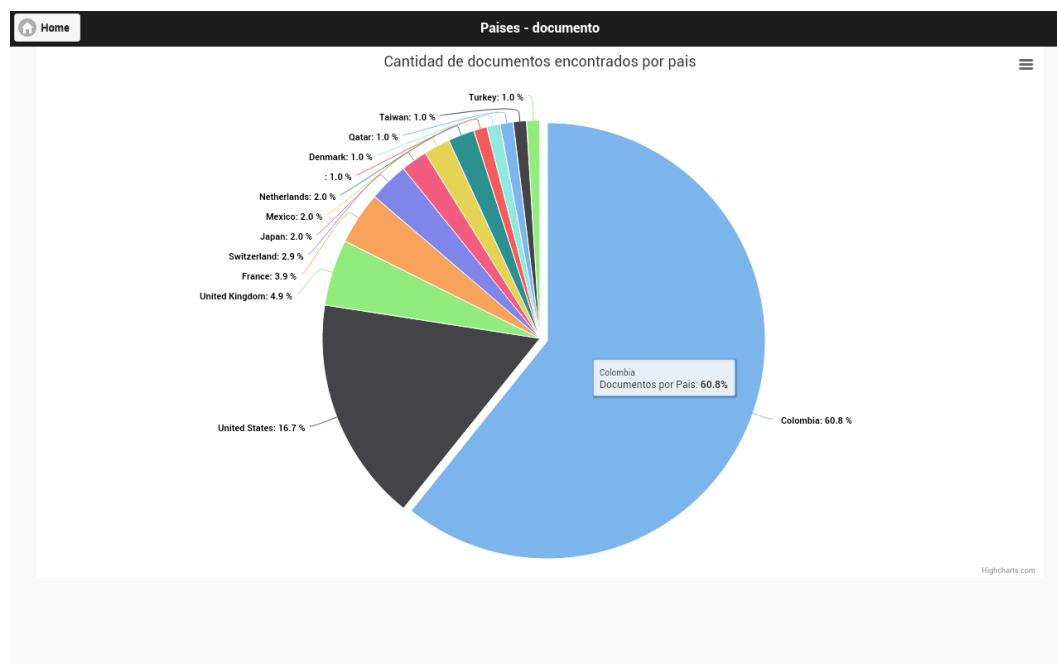


Figura 11 Circular Países - documento

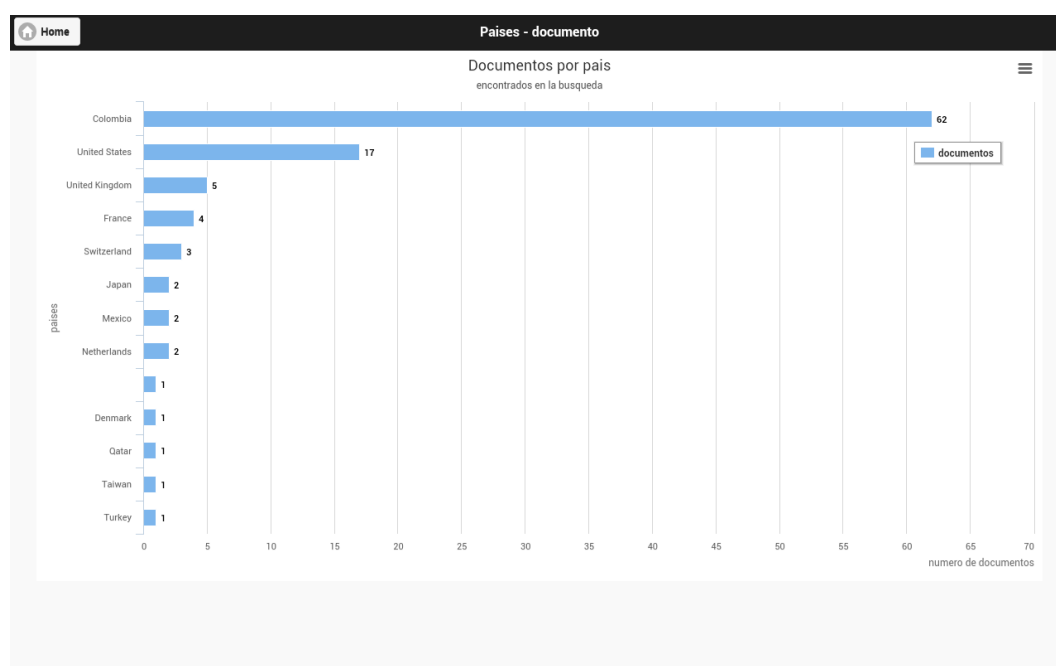


Figura 12 Barras Países - documento

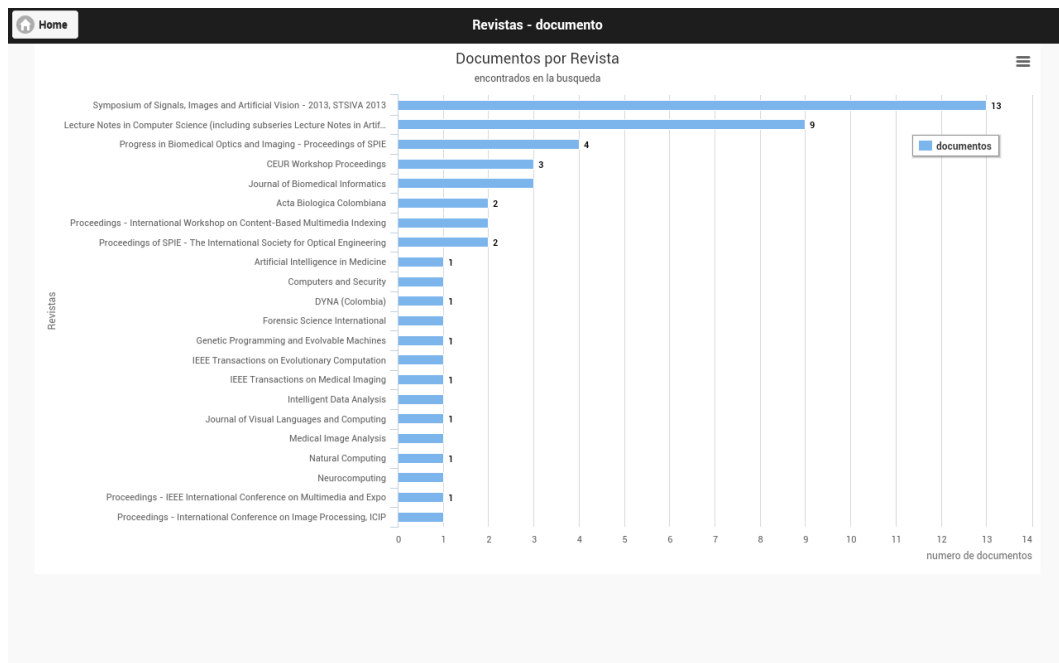


Figura 13 Barras Revistas - documento

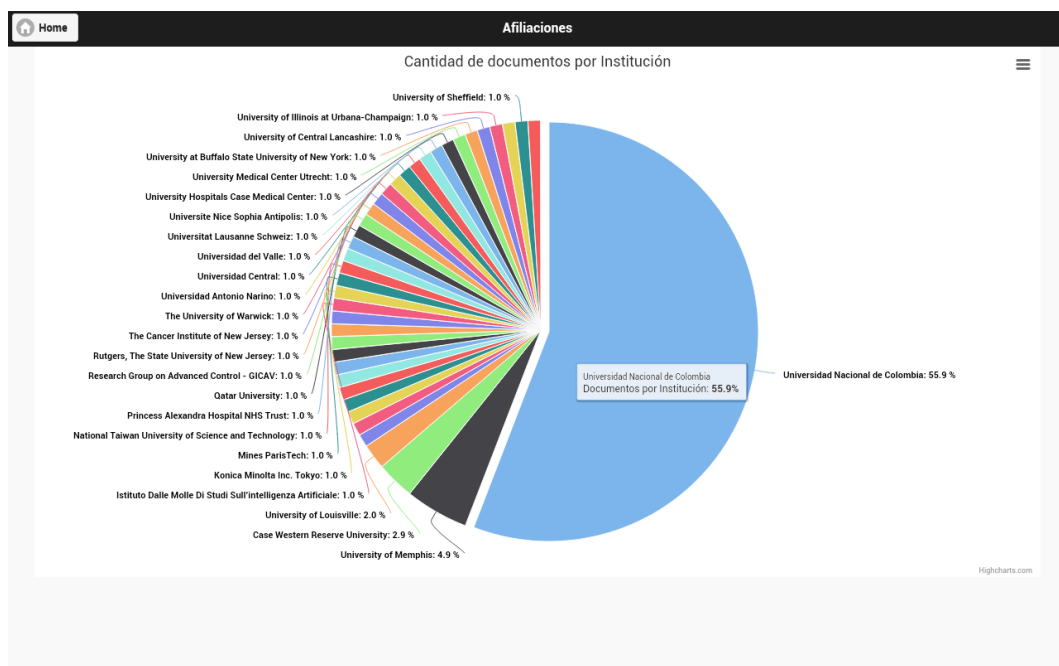


Figura 14 Circular Instituciones - documento (Afiliaciones de los autores)

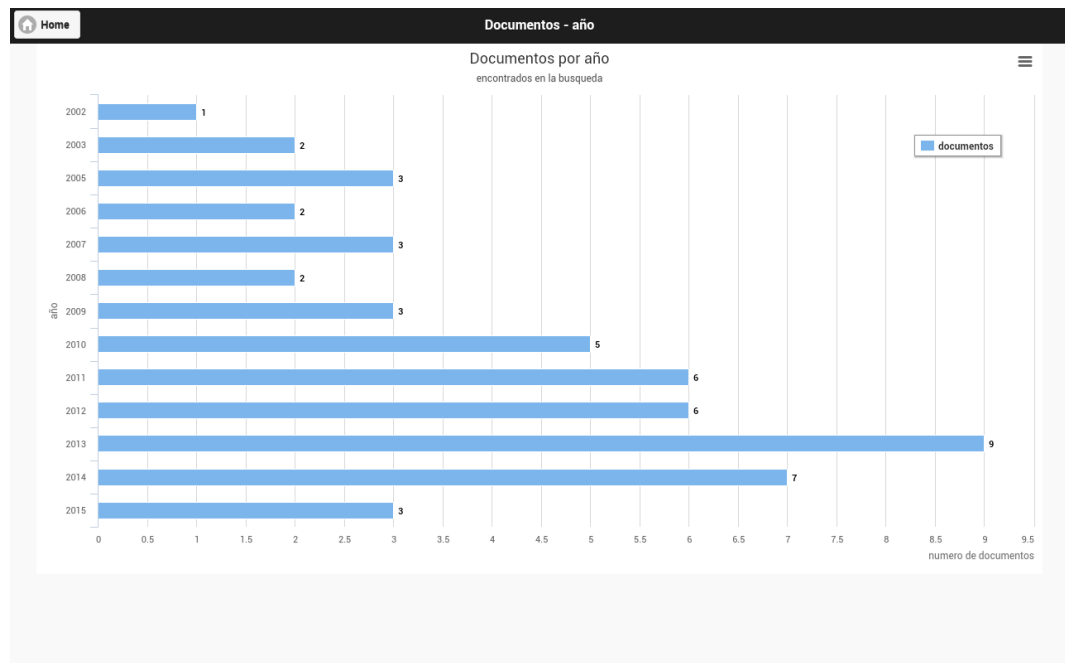


Figura 15 Barras Documentos por año

En tipos de análisis se encuentra clustering el cual corresponde al agrupamiento de documentos por características similares, cada uno de los documentos queda en un grupo particular que corresponda con sus características y es visualizado en VT-Mobile desde el menú de analíticas como una nube de figuras circulares como se muestra más adelante en la figura 16.

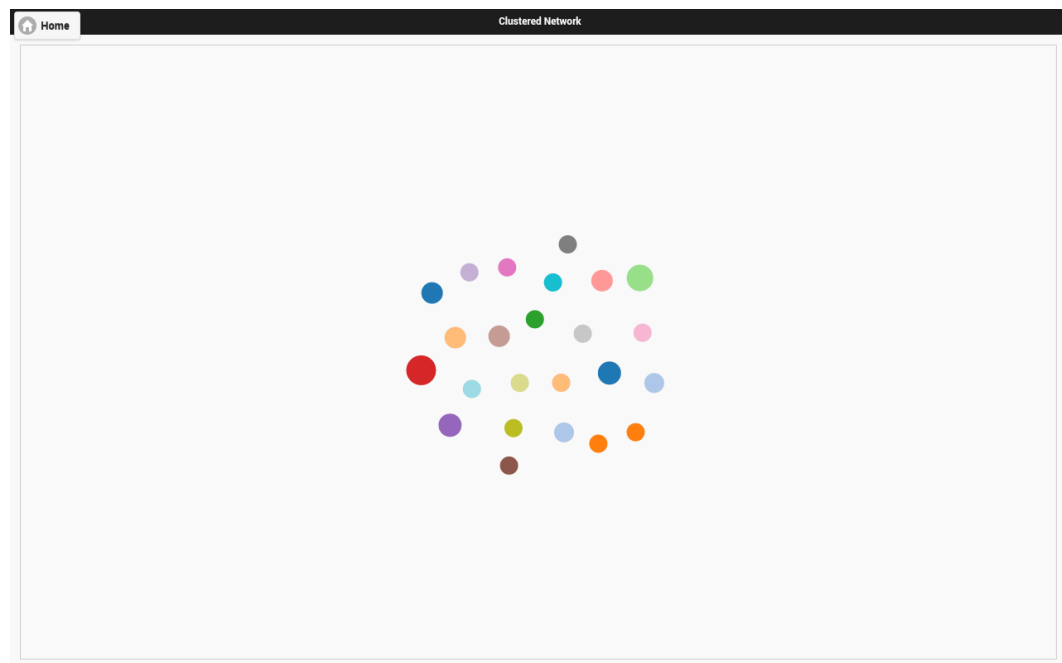


Figura 16 Clustering

La figura 16 muestra una nube de puntos que corresponde a cada uno de los grupos en los cuales fueron agrupados los documentos, entre mayor sea la cantidad de documentos dentro de un grupo en particular, mayor será el tamaño de este grupo dentro de la visualización.

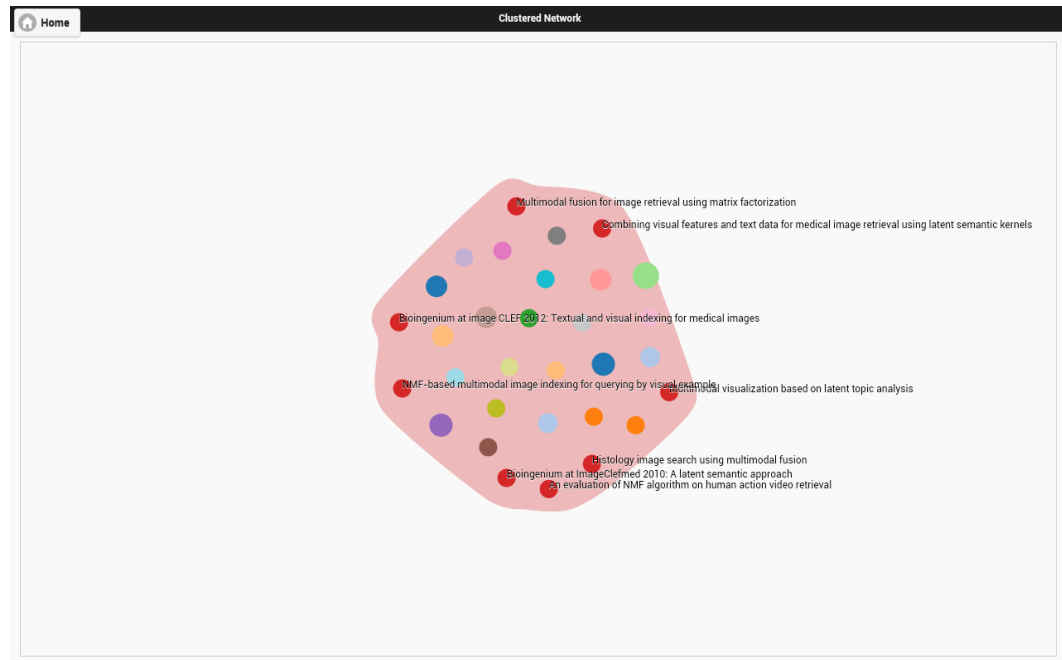


Figura 17 Clustering expansión del grupo más grande

Al tocar cualquiera de los grupos se obtiene lo que se muestra en la figura 17, el grupo se ha dividido en elementos más pequeños que corresponden a cada uno de los documentos dentro de ese grupo debidamente etiquetados por el nombre del documento. En las figuras siguientes se muestran otros grupos.

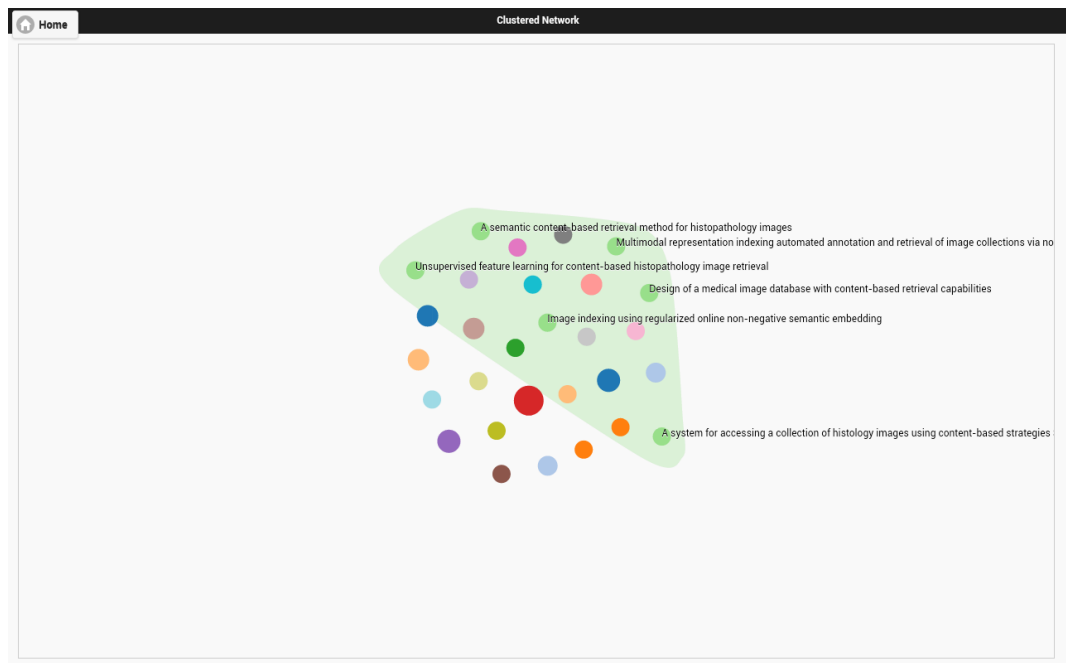


Figura 18 Clustering expansión segundo grupo

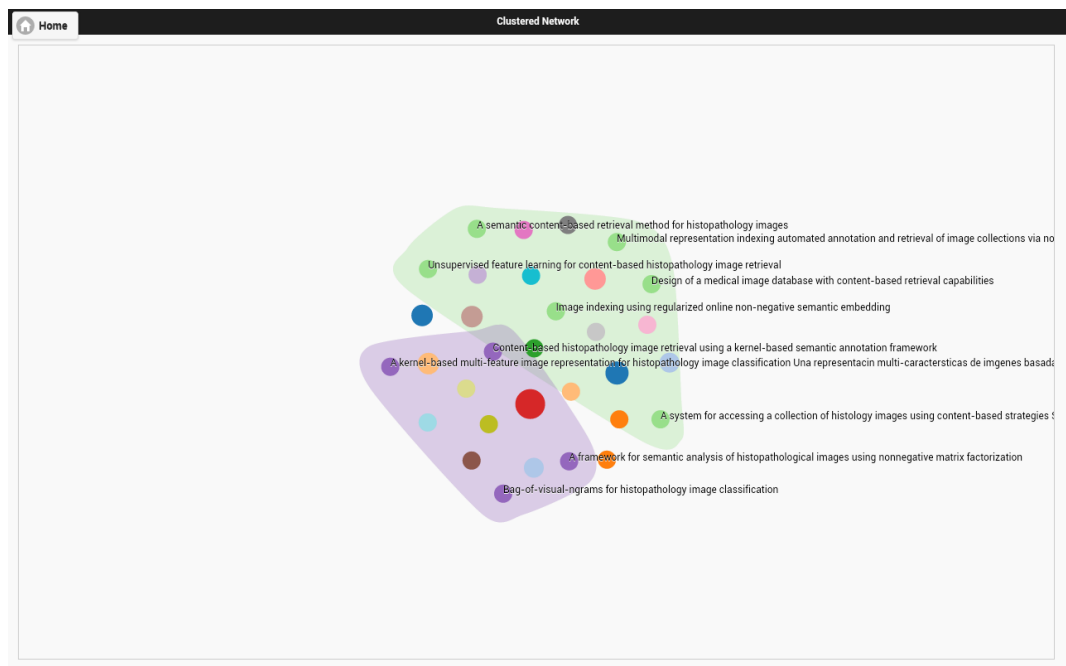


Figura 19 Clustering expansión tercer grupo

De esta forma se visualizan los resultados del agrupamiento correspondientes a los documentos del autor consultado, finalmente se muestra otra opción de VT-Mobile que permite acceder a los documentos encontrados en Scopus, como se muestra en la figura 20.

← Atras Documentos	
3-Functions meta-learner algorithm: A mixture of experts technique to improve regression models	➤
A comparative analysis of artificial immune network models	➤
A comparative evaluation of supervised and unsupervised representation learning approaches for anaplastic medulloblastoma differentiation	➤
A framework for high performance image analysis pipelines	➤
A framework for semantic analysis of histopathological images using nonnegative matrix factorization	➤
A grid computing approach to subtraction radiography	➤
A kernel-based framework for image collection exploration	➤
A kernel-based multi-feature image representation for histopathology image classification Una representación multi-características de imágenes basada en kernels par...	➤
A multi-class kernel alignment method for image collection summarization	➤
A neuro-immune model for discriminating and visualizing anomalies	➤
A scalable artificial immune system model for dynamic unsupervised learning	➤
A semantic content-based retrieval method for histopathology images	➤
A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images	➤
A system for accessing a collection of histology images using content-based strategies Sistema para acceder una colección de imágenes histológicas mediante estrate...	➤
An adaptive image representation learned from data for Cervix cancer tumor detection	➤
An evaluation of NMF algorithm on human action video retrieval	➤
An immunity-based technique to characterize intrusions in computer networks	➤
Anomaly detection using real-valued negative selection	➤

Figura 20 Lista de documentos

3.1.2 Conclusiones de la prueba piloto

se realizó una prueba piloto con un autor en específico que hace parte de una comunidad académica y que tiene una producción científica notable, se revisaron todos los datos que la aplicación presento como respuesta, se presentó todo el proceso y se explicó lo que VT-Mobile puede hacer, se realizó también la misma búsqueda en Scopus donde los resultados fueron similares, Scopus realiza sus propias visualizaciones sobre la búsqueda y muestra indicadores, los cuales se asemejaban a los mostrados por la aplicación, además cada uno de los documentos encontrados por la aplicación pertenecían a documentos relacionados con este autor en Scopus por lo tanto la información que despliega VT-Mobile si corresponde con la producción científica del autor percibida por Scopus y permite hacer búsquedas y análisis desde dispositivos móviles de manera sencilla y rápida, poniendo esta información y análisis propios a disposición del usuario para que tome decisiones o cree opiniones acerca de un tema o autor en particular.

3.2 Prueba para la experiencia de usuario

La aplicación fue probada por los docentes de la EISC a través de un ejercicio sencillo el cual será descrito a continuación y al final se realizó una encuesta para obtener un concepto de su percepción de la aplicación respecto a Scopus la fuente de los datos enfatizando en sus capacidades móviles, sencillez y capacidades adicionales la prueba se describe a continuación.

3.2.1.1 Ejercicio

El ejercicio planteado para dar a conocer la aplicación a los profesores fue realizar una entrevista corta con los profesores y proponiendo 2 búsquedas de la producción científica de un docente específico la primera búsqueda en Scopus para mostrar los resultados y los análisis que Scopus presenta sobre los documentos encontrados a continuación realizar la búsqueda del mismo docente desde VT-Mobile y que el sujeto del ejercicio navegue por las opciones que la aplicación le presenta para desplegar la información. Para el ejercicio se realizaron búsquedas de la producción científica del encuestado u otro profesor de la escuela que tuviera publicaciones en común.

3.2.1.2 Encuesta

La primera pregunta corresponde al nombre del docente:

Alexander Aranda	29/05/2015 17:28	Vea las respuestas del encuestado
John Sanabria	29/05/2015 12:07	Vea las respuestas del encuestado
Angel Garcia Baños	29/05/2015 11:59	Vea las respuestas del encuestado
Liliana Machuca	29/05/2015 11:47	Vea las respuestas del encuestado
pedro moreno	29/05/2015 10:46	Vea las respuestas del encuestado
IRENE TISCHER	28/05/2015 12:28	Vea las respuestas del encuestado
Victor Manuel Vargas Forero	27/05/2015 18:55	Vea las respuestas del encuestado
Daniel Barragán	27/05/2015 18:20	Vea las respuestas del encuestado
Oscar Ceballos	27/05/2015 18:01	Vea las respuestas del encuestado

Figura 21 Lista de docentes

En total se realizó la prueba con 9 docentes de la EISC.

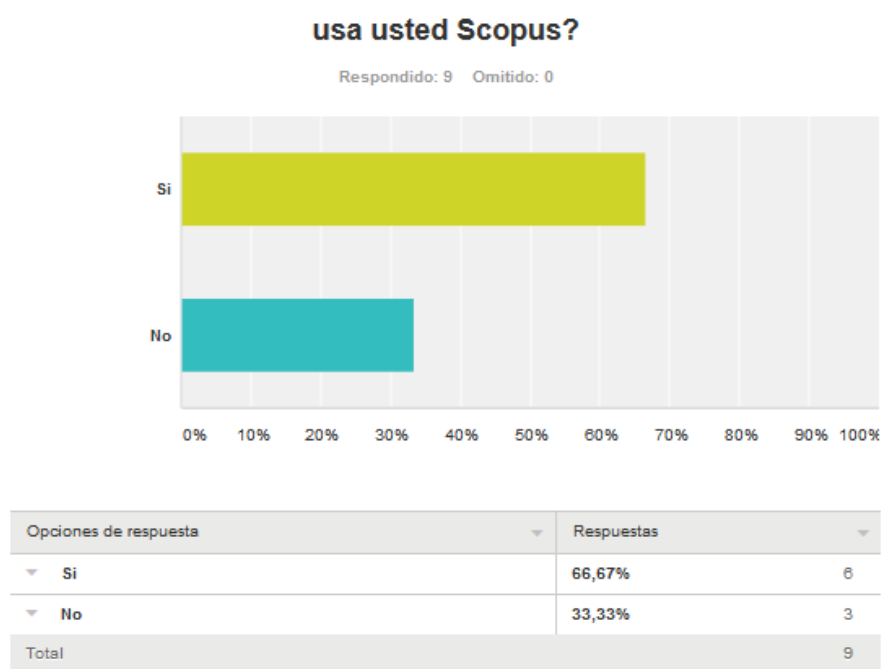


Figura 22 Pregunta 2

Algunos docentes no estaban muy familiarizados con Scopus.

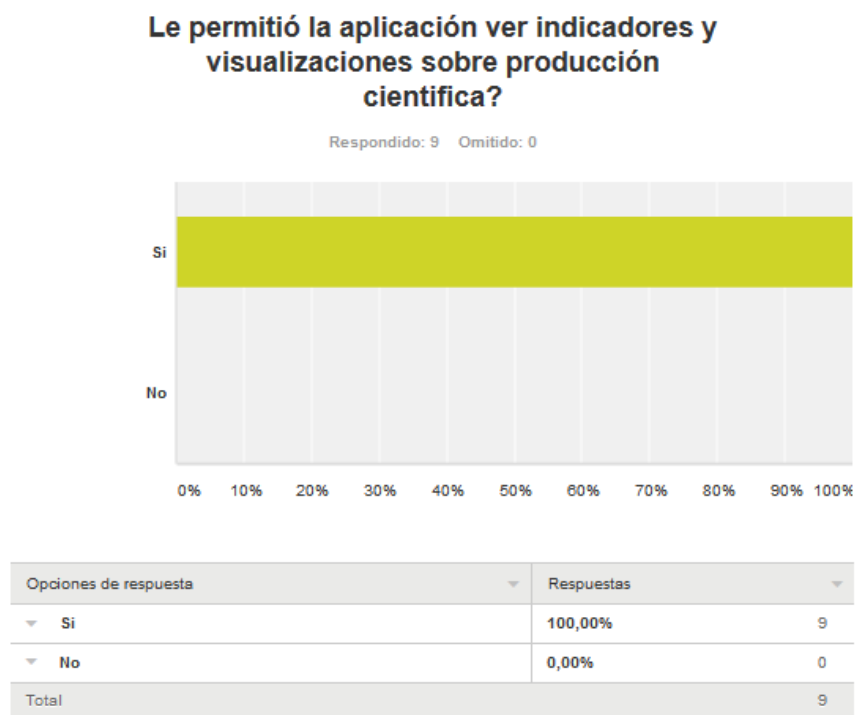
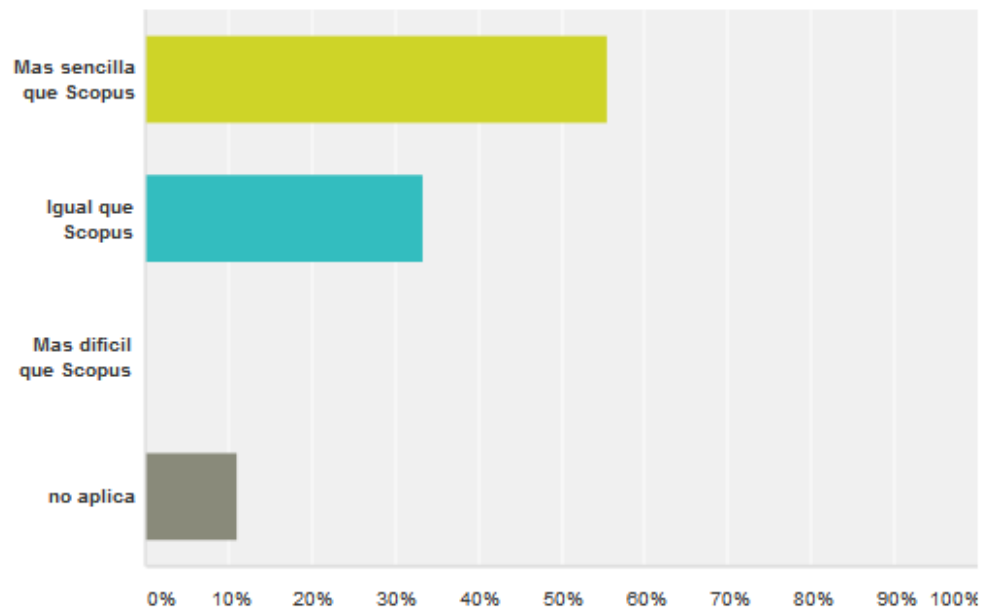


Figura 23 Pregunta 3

Su experiencia con la aplicación fue más sencilla o igual que su experiencia con Scopus?

Respondido: 9 Omitido: 0



Opciones de respuesta	Respuestas	
Mas sencilla que Scopus	55,56%	5
Igual que Scopus	33,33%	3
Mas dificil que Scopus	0,00%	0
no aplica	11,11%	1
Total		9

Figura 24 Pregunta 4

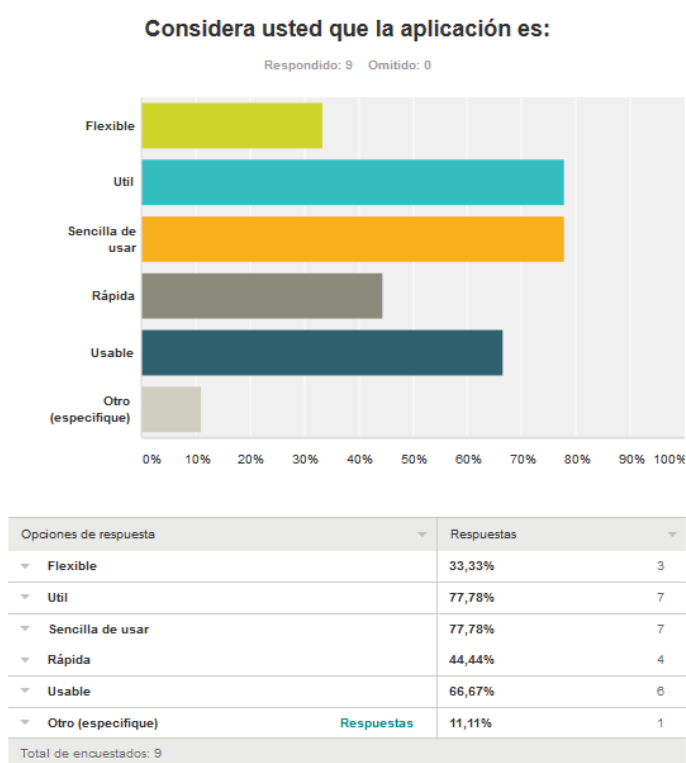


Figura 25 Pregunta 5

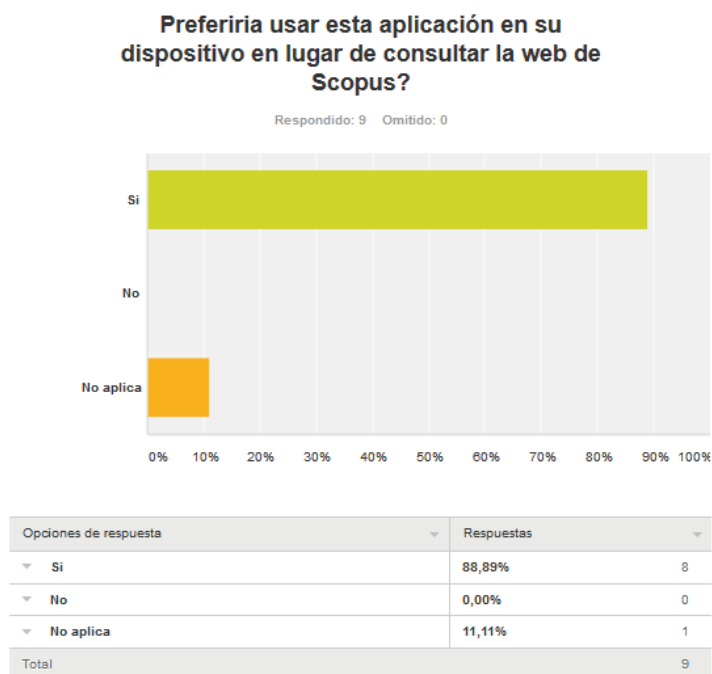


Figura 26 Pregunta 6

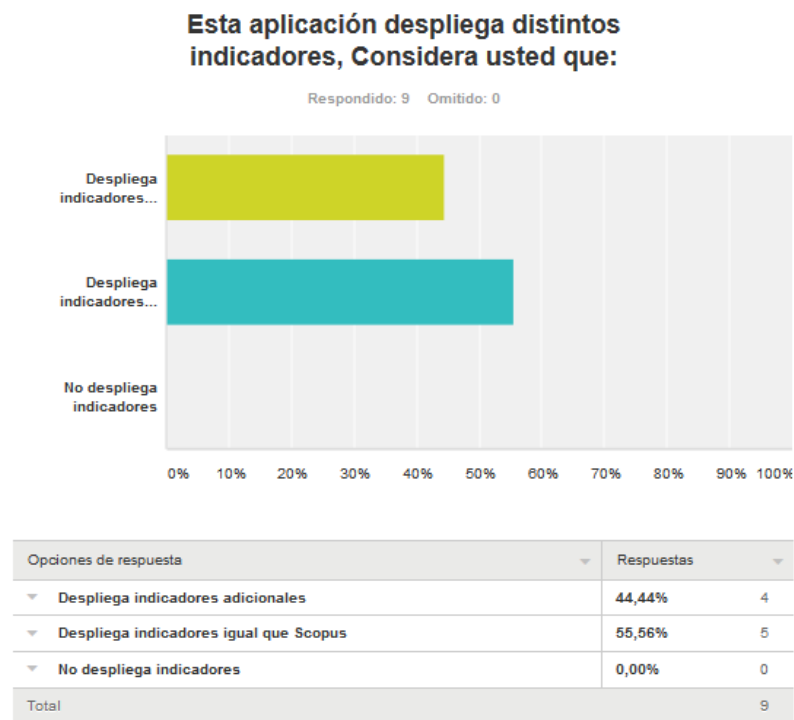


Figura 27 Pregunta 7

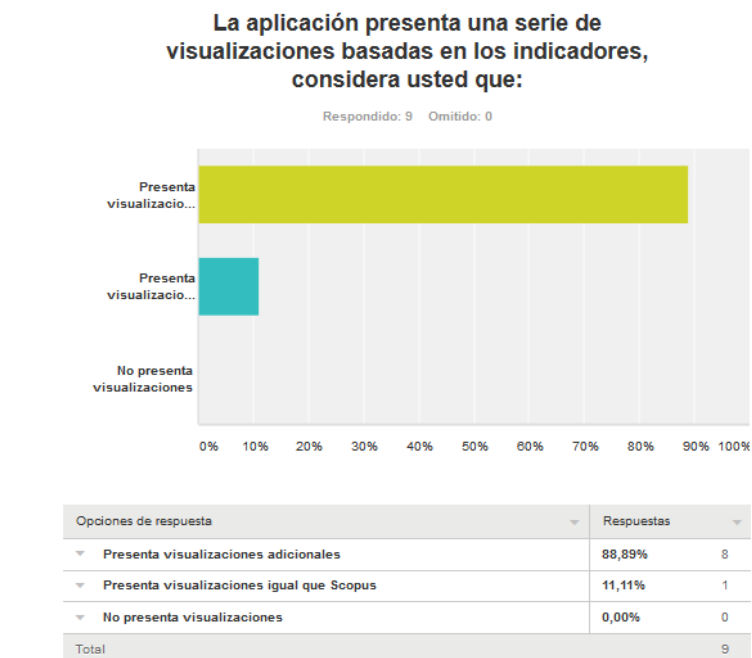
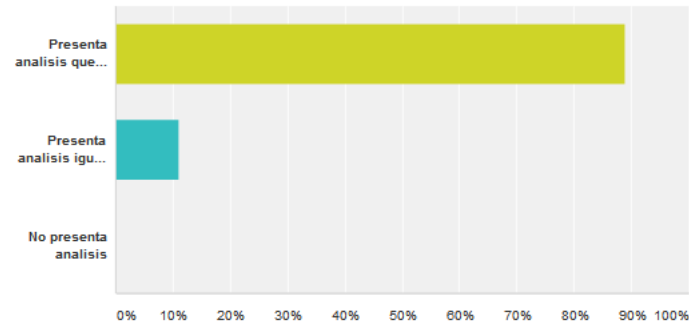


Figura 28 Pregunta 8

La aplicación presenta analisis (ej. Clustering), considera usted que:

Respondido: 9 Omitido: 0

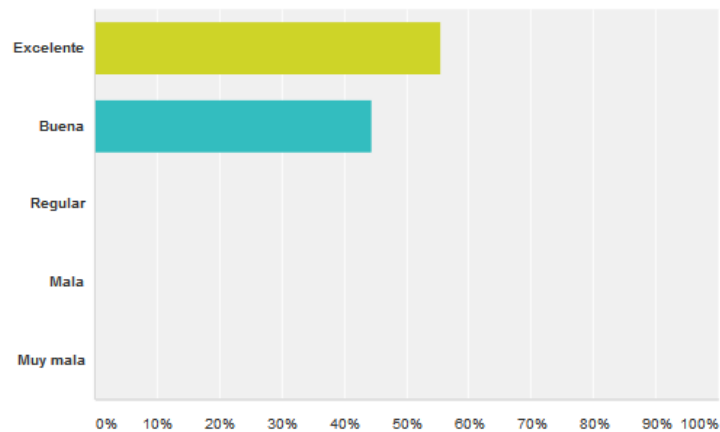


Opciones de respuesta	Respuestas	
Presenta analisis que no estan en Scopus.	88,89%	8
Presenta analisis igual que Scopus.	11,11%	1
No presenta analisis	0,00%	0
Total		9

Figura 29 Pregunta 9

Califique su experiencia con la herramienta

Respondido: 9 Omitido: 0



Opciones de respuesta	Respuestas	
Excelente	55,56%	5
Buena	44,44%	4
Regular	0,00%	0
Mala	0,00%	0
Muy mala	0,00%	0
Total		9

Figura 30 Pregunta 10

3.2.2 Conclusiones prueba de experiencia de usuario

Después de realizar el ejercicio y analizar los resultados de la encuesta realizada se concluye que la mayoría de los docentes participantes consideran que VT-Mobile tiene características adicionales al análisis básico de Scopus y también que la aplicación es bastante sencilla de usar, durante la prueba se notó como algunos de los participantes se interesaban más por sus publicaciones y citas, sorprendiéndose incluso con publicaciones citadas que ya no recordaban se reconoce a la herramienta como una aplicación útil, capaz de analizar la producción científica de los docentes de la escuela y desplegar visualizaciones. Esta prueba confirma el impacto positivo que puede generar la aplicación dentro de la EISC y la buena disposición de los docentes frente a la posibilidad de uso de esta herramienta, su integración con la plataforma Vigtech Univalle y la mejora posterior de la aplicación.

3.3 Conclusiones

Se implementó una herramienta móvil, con la capacidad de buscar, extraer, analizar y visualizar, datos encontrados en Scopus para el beneficio de la comunidad académica estudiada y las prácticas de vigilancia científico-tecnológica, ofreciendo una manera simple de revisar la producción científica de los docentes dentro de esta comunidad y apoyarse en esto para tomar decisiones o dar conceptos acerca de las prácticas actuales dirigidas a la creación de nuevo conocimiento dentro de la Universidad del Valle. La aplicación permite buscar y analizar información desde una plataforma móvil sin tener que acceder directamente al sitio de Scopus, de una forma simple y guardando los resultados de la búsqueda anterior para eliminar la necesidad de una conexión constante, además se realiza y despliega una visualización de un análisis de Clustering que no se puede obtener desde Scopus. La aplicación está fundamentada en procesos de la vigilancia tecnológica junto a todas las técnicas y conceptos usados para lograr cumplir con estos procesos. Se basa en Vigtech una aplicación robusta y la plataforma Vigtech Univalle de la cual formara parte, usando diferentes tecnologías ya conocidas para su implementación. En resumidas palabras ahora existe una herramienta para el apoyo de procesos de vigilancia científico-tecnológica dentro de comunidades académicas.

3.4 Trabajo Futuro

Para el trabajo futuro se plantea la mejora del aspecto visual de VT-Mobile junto a la integración con la plataforma Vigtech Univalle añadiendo nuevos análisis a la herramienta. Ampliar el caso de estudio y realizar pruebas con docentes o investigadores por fuera de la EISC. En otras palabras hacer de VT-Mobile una herramienta sencilla de usar, agradable para el usuario con una integración a una plataforma robusta diseñada para apoyar la vigilancia tecnológica en la Universidad del Valle.

REFERENCIAS

- Sánchez Torres, J. M., & Palop Marro, F. (2008). Herramientas de Software especializadas para Vigilancia Tecnológica e Inteligencia Competitiva., 1–15.
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. En Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 249–256). ACM. Recuperado a partir de <http://dl.acm.org/citation.cfm?id=97281>
- Quintero, C. A., Florian-Gaviria, B., & Pabon, O. S. (2014, September). Comparative study of technologies for mobile learning analytics. In Computing Colombian Conference (9CCC), 2014 9th (pp. 82-89). IEEE
- López, C. P. (2007). *Minería de datos: técnicas y herramientas*. Editorial Paraninfo.
- Bucheli, V. A. (2007). Herramienta informática para vigilancia tecnológica VIGTECH Technology Monitoring Software tool VIGTECH. *Revista Avances en Sistemas e Informática*, 4(1), 118.
- Gentleman, R., Ihaka, R., & Bates, D. (2009). The R project for statistical computing. URL: <http://www.r-project.org/254>.
- Sutherland, J., & Schwaber, K. (2007). The Scrum Papers. Scrum. Recuperado a partir de http://scruminc.com/tl_files/scrum_inc/documents/ScrumPapers.pdf
- Anthes, G. (2011). Invasion of the Mobile Apps. *Commun. ACM*, 54(9), 16–18. doi:10.1145/1995376.1995383
- Gould, J. D., & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM*, 28(3), 300-311. doi:10.1145/3166.3170
- Cockton, G. (2008). Revisiting usability's three key principles (p. 2473). ACM Press. doi:10.1145/1358628.1358704
- Chittaro, L. (2006). Visualizing information on mobile devices. *Computer*, 39(3), 40-45. doi:10.1109/MC.2006.109
- Fowler, M., & Highsmith, J. (2001). The agile manifesto. *Software Development*, 9(8), 28–35.
- Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Ramírez, M. I., Rua, D. E., & Alzate, B. A. (2012). Vigilancia tecnológica e inteligencia competitiva. *Gestión de las Personas y Tecnología*, (13), 149-153.